Writing a Data Management Plan
Why Manage Your Data?

Posting to widely-searched repositories increases use and impact

Ensures longer-term preservation

Increases efficiency by standardizing data
Why Manage Your Data?

Ensures Data Integrity

Requirement for Funders

Prevents Duplication

Supports Open Access
Data Life Cycle
Data Format

- Text
- Numerical
- Multimedia
- Models
- Software
- Hardware Specific
Basic Questions

What type of data will be produced?
How much?
How fast will it grow?
How often will it change?
Who is the audience?
Who owns the data?
Who is responsible for the data?
How long should it be retained?
Basic Questions

What type of data will be produced?
How much?
How fast will it grow?
How often will it change?
Who is the audience?
Who owns the data?
Who is responsible for the data?
How long should it be retained?
Components of a Plan

- Project Abstract
- Nature and Format of Data
- How Will Data Be Collected
- Format Standard
- Metadata Standard
Short-Term Section

File Formats
Local Storage Policies
Back-Up Policies
Security Policy
List of Responsible Groups with Contact Info
Legal and Ethical Issues Section

- Intellectual Property
- Any Human Subject Issues
- Itemized List of Any Data Limitations
- Listing of all Associated Legal Documentation
Access Policies

How is the data made available?
(github, data dumps, etc)

Data Restrictions
Use Restrictions
Long-Term Preservation

Will all data be preserved or just the final products?
Use restrictions
Schedule of any conversions or forensic checks
Staffing

Who is responsible for each data management task?

Where do the handoffs occur?
Cost

Estimate the cost of your data plan including:

- Storage
- Staffing
- Conversion/Forensics
Data Integration

- eXtensible Markup Language (XML)
- Resource Description Framework (RDF)
- Uniform Resource Identifiers (URIs)
- Ontologies (structured reference vocabs)
Online Tools

DMPTool: https://dmp.cdlib.org/
DMPOnline: https://dmponline.dcc.ac.uk/